OLASILIK TEORİSİ VE İSTATİSTİK

Bağımsız Olaylar

- Ortaya çıkmaları birbirlerine bağlı olmayan olaylardır.
- Örneğin iki zar birlikte atılsa, zarlardan birinin 3 diğerinin 2 gelmesi olasılığı bağımsız iki olaydır.
- İki ya da daha çok olayın ortaya çıkması birbirlerine bağlı değilse bu olaylara bağımsız olaylar diyoruz.
- A ve B bağımsız iki olay ise, bu olaylar için $P(A/B)=P(A)$ ve $P(B/A)=P(B)$ dir.

- A ve B olaylarının bağımsız olması için gerek ve yeter koşul:

$$P(A \cap B) = P(A).P(B)$$

- Bunu kanıtlamak için koşullu olasılık formülünden

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

- yazılabilir. Her iki kesrin payı aynı olduğu için

$$P(A \cap B) = P(A/B)xP(B) = P(B/A)xP(A)$$

$$P(A/B) = P(A)$$

$$P(B/A) = P(B)$$

- eşitliğinden

$$P(A \cap B) = P(A).P(B)$$

- **Problem 1:** Bir para ve bir altı yüzlü zar birlikte atılıyor.

  $A:\{Parada \quad T\}$, $B:\{Zarda \quad 3 \quad gelme\}$ olarak tanımlansın. A ve B olayları bağımsız mıdır?

  $S:\{T1, T2, ..., T6, Y1, Y2, ..., Y6\}$

  $P(A \cap B) = \frac{1}{12} \quad P(A) = \frac{1}{2} \quad P(B) = \frac{1}{6}$

  $P(A \cap B) = P(A).P(B)$

  $\frac{1}{12} = \frac{1}{2} \frac{1}{6}$

eşitliği sağlandığı için A ve B olayları bağımsızdır.

- **Problem 2:** Ekonomi sayfasını okuma araştırmasında [(26-40) yaş arası olma], (Ara sıra okuma) olayları bağımsız mıdır?

$$P\left[(26-40)\ yaş \quad arası \quad olma\right]=0,36$$

$$P\left(Arasıra \quad okuma\right)=0,45$$

$$P\left[(26-40)\ yaş \cap \quad Arasıra \quad okuma\right]=0,13$$

$$0,36x0,45=0,16$$

olup, görüldüğü gibi 0.13≠0.16 olup,bağımsız olaylar değildir.

BAYES Teoremi

- Bayes Teoremine geçmeden Toplam olasılık formülünü vermemiz yararlı olur.
- Toplam olasılık formülünde S örneklem uzayının B1 , B2 ... Bk gibi birbirinden ayrık ve Toplamları S örneklem uzayını oluşturacak alt uzaylara ayrıldığı düşünülür.
- **Örnek 1:** Bir sigorta şirketi bir bölgede sağlık sigortaları için faaliyet gösterecektir. Ön araştırma için o bölgede yaşayan kitlenin yaş dağılımı ile ilgili bir araştırma yapmıştır.

Yaşlar	0-15	16-25	26-40	41-55	56→
GöreliSıklık (olasılık)	0,16	0,25	0,22	0,25	0,12

- Yaşlar ile ilgili dağılımı Çizelge de verildiği gibi bulmuştur.
- Bu çizelgede olay 0-56→ yaşlarıdır.
- $E_1:\{0-15\}$, $E_2:\{16-25\}$, $E_3:\{26-40\}$, $E_4:\{41-55\}$, $E_5:\{56\rightarrow\}$ olan yaş gruplarıdır. Bu yaş grupları ayrık ve toplamları S yi verir.
- Bu bölgede bir anket uygulandığını ve ankette, A olayını sağlık sigortası yaptırmayı düşünenler olarak tanımlayalım.

$$P(A) = P\left[(A \cap E_1) \cup (A \cap E_2) \cup (A \cap E_3) \cup (A \cap E_4) \cup (A \cap E_5)\right]$$

- $A \cap E_i$ **'ler ayrık olduğu için** $P(A) = P(A \cap E_1) + P(A \cap E_2) + ... + P(A \cap E_5)$

yazılır. Buradan;

$$P(A) = \sum_{i=1}^{6} P(A \cap E_i)$$

- **Koşullu olasılık verilen formülden**

P(A∩Ei) = $P(E_i)P(A/E_i)$

olarak yazılabilir. Bu durumda $P(A) = \sum_{i=1}^{6} P(E_i)P(A/E_i)$

elde edilir. Ei'lerin n tane olduğunu düşünüp bu formülü $P(A) = \sum_{i=1}^{n} P(E_i)P(A/E_i)$

yazabiliriz. Bu formül toplam olasılık formülü adını alır. Yukarıdaki örnekte yapılan araştırmada;

$$P(A/E_1)=0,01 \quad P(A/E_2)=0,10 \quad P(A/E_3)=0,09 \quad P(A/E_4)=0,12 \quad P(A/E_5)=0,20$$

- olduğu belirlenmiş olsun. P(A) bu bölge için ne kadardır. Toplam olasılık formülü kullanılarak;

P(A)=0.16x0.01+0.25x0.10+0.22x0.09+0.25 x0.12+0.12x0.20=0.1004

- olarak bulunur.

Bayes Formülü:

- Bu formülü yukarıda verdiğimiz örnek ile açıklamaya çalışalım.
- Bu bölgeden herhangi bir kişi belirlenmiş ve sağlık sigortasına baş vurmayı düşündüğü anlaşılmıştır.
- Bu kişinin hangi yaş grubundan olabileceğini araştıralım.
- Burada koşul sağlık sigortası yaptırmayı düşünmüş olmasıdır. Yani;

- $P(E_i / A) = \dfrac{P(E_i \cap A)}{P(A)}$ dır.
- $P(E_i \cap A) = P(E_i)P(A/E_i)$ yazılabileceğini

- Bölüm8.3'de gördük. P(A) içinde toplam olasılık formülünden;

$$P(A) = P(E_1)P(A/E_1) + P(E_2)P(A/E_2) + P(E_3)P(A/E_3) + P(E_4)P(A/E_4) + P(E_5)P(A/E_5)$$

olur

- P(Ei/A) eşitliğinde bu bilgiler genellenerek yerlerine konulur ise

- $$P\left(E_i / A\right) = \frac{P\left(E_i\right) P\left(A / E_i\right)}{\sum_{i=1}^{n} P\left(E_i\right) P\left(A / E_i\right)}$$ elde edilir.

- Bu formüle Bayes Formülü denilmektedir.

- Bu eşitlikte, $P(E_i)$ $i=1, 2, ..., n$ olasılıklarına önsel olasılıklar denir. Bu olasılıklar konu ile ilgili verilebilecek önceki olasılıklardır.

- $P(E_i/A)$ $i=1, ..., n$ olasılıklarına da sonsal olasılıklar (ardıl) denir.

- **Örnek 1.** Sigorta şirketi örneğinde; Bir kişi sigorta yaptırmayı düşündüğüne göre 56 ve daha ileri yaş grubunda olma olasılığını bulalım.
- **Çözüm:** $P(56 \rightarrow / A) = \dfrac{P[(56 \rightarrow) \cap A]}{P(A)}$

$$E_5$$

$$P\left[(56 \rightarrow) \cap A\right] = P(E_5) \times P(A / E_5)$$

= 0,12 x 0,20 = 0,024 P(A) = 0,1004 olarak hesaplandı

$$P(E_5 / A) = \frac{0.024}{0.1004} = 0.239$$

$P(E_1 / A)$, $P(E_2 / A)$, $P(E_3 / A)$, $P(E_4 / A)$ bulalım

- **Örnek 2.** Bir şirketin hesap bakiyelerinin eski kayıtlarını inceleyen bir denetmen hesapların %15'nin hata içerdiğini bulmuştur. Bu hatalı bakiyelerin %60'ı eski rakamlara göre olağan dışı değerler (olması mümkün olmayan sonuçlar) olarak görülmektedir. Bütün hesap bakiyelerinde %20'si olağan dışı değerlerdir. Belli bir bakiye olağan dışı gibi görülse hatalı çıkma olasılığı nedir?

  İlgilendiğimiz olaya (hata) diyelim.
- P(hata) = 0,15   P(olağandışı değer) = 0,20
- P(olağandışı değer / hata) =0,60

  Bayes teoreminden

$$P\left(hata / olağandışı \quad değer\right) = \frac{P\left(olağandışı \quad değer / hata\right) \times P\left(hata\right)}{P\left(olağandışı \quad değer\right)}$$

$$= \frac{0,60 \times 0,15}{0,20} = 0,45$$

Rasgele Değişkenler

- Olasılık, kitle ile örneklem arasında olan bağıntıyı kurmamıza yardımcı olacaktır. Bu amaçla önceki konularımızla, raslantıya bağlı olaylardan ve bu olayların gerçekleşme olasılıklarından söz ettik. Konuyu daha kolay kavrayabilmek için de küçük olaylardan örnekler verdik.
- Örneğin paranın atılıp yazı ya da tura gelmesi, iskambil kağıtlarından kart çekilmesi, altı yüzlü zar deneyi gibi. Bu deneyler birer rasgele deneydir.
- Rasgele deneylerin tüm gerçekleşebilecek sonuçlarını da örneklem uzayı (S), bu sonuçlar içinde tanımlanan özel sonuçları oluşturduğu alt kümeye de olay dedik.
- Genellikle de A, B, C harfleri ile gösterdik. Bu konuda yeni bir tanım vereceğiz. Rasgele deney sonucunda tanımlanan bir özelliği sayısal bir değerle ifade edebiliyorsak, bu özellik rasgele değişken olarak isimlendirilecektir.

- **Rasgele Değişken:** Değeri sayısal olarak rasgele deney sonucu belirlenen değişkendir.

- Örnekler:
- Bir basketbol takımının bir sezon boyunca her maçta ayrı ayrı kazandığı sayılar
- İki paranın birlikte atılması deneyinde gelen yüzlerde yazı gelme sayısı
- Bir metro istasyonunda belirlenen, bir ay içinde belli bir zaman aralığında (sabah 9 ile 12 arasında) gelen müşterilerin günlük sayıları
- Bir paranın yazı gelinceye kadar atılma sayıları
- Bir denetçinin hatalı dosya buluncaya kadar inceleyeceği dosya sayısı
- Bir gıda maddesinde koruyucu madde oranı
- Paketlenmiş olarak satılan bir ürünün ağırlığı

- Örneklere yeniden dönelim. Yukarıda verilen birinci, ikinci üçüncü örneklerde raslantı değişkenlerinin sayısal değerleri sayı olarak belirlenmektedir.
- Birinci örnekte bir sezonda 34 maç olup, birinci maçta 82, ikinci maçta 70 ve...34. maçta 65 sayı yapılmış olabilir. Burada örneklem uzayı;

  S:{82,70,...,65} olacaktır.

  İkinci örnekte örneklem uzayı;

  S:{YY,YT,TY,TT},

  üçüncü örnekte örneklem uzayı;

  S:{102,90,95,...,110}
- olabilir. Bu üç örnekte raslantı değişkeni sayılabilir sonlu özelliktedir.

- Dördüncü ve beşinci örneklerde rasgele değişken yine sayılabilir ancak nerede duracağımız belli değildir.
- Örneğin dördüncü örnekte paranın yazı gelinceye kadar atılmasında örneklem uzayı;
- S:{Y,TY,TTY,TTTY,...}
- olacaktır.
- Beşinci örnekte de durum benzerdir. Denetçi İlk dosyada da hata bulabilir, ikincide de...30. dosyada da ve böylece devam edebilir. Bu örneklerde rasgele değişken sayılabilir sonsuzluktadır.

- Altıncı örnekte verilen rasgele değişken katkı maddesinin bulunma oranı olup %0.1, %0.12... gibi değerler alabilir.
- Yedinci örnekte paketlerin ağırlıkları çok hassas ölçülürse belli iki değer arasında her ağırlığı alması mümkündür.
- Bu iki örnekte raslantı değişkeni süreklidir. Bu örneklerden sonra raslantı değişkenleri; sonlu ya da sonsuz sayılabilir özellikte ise kesikli raslantı değişkeni, bir ya da daha çok aralıkta her değeri alabiliyorsa, sürekli raslantı değişkeni olarak anılacaktır.
- Raslantı değişkenleri X, Y, Z gibi büyük harflerle, bu değişkenlerin aldığı sayısal değerler de x, y, z gibi küçük harflerle gösterilecektir. Örneğin yukarıda verilen birinci örnekte sezon boyunca her maçta kazanılan sayılar X, alınan değerler ise x_1=82, x_2=70,...,x_{34}=65 olacaktır.

Kesikli Rasgele Değişkenin Dağılımı

Bir raslantı değişkeninin alabileceği değerlerin o değeri alma olasılıkları ile birlikte belirtilmesine olasılık dağılımı ya da olasılık fonksiyonu denir.

- **Örnek:** Önce verdiğimiz paranın iki kez atılma deneyini düşünelim.
- $S:\{YT, TY, TT, YY\}$ idi.
- X tura gelme saysı olsun.
- x_1=0(hiç tura gelmeyebilir),
- x_2=1(bir kez tura gelebilir),
- x_3=2(iki kez tura gelebilir).
- Burada olasılık dağılımını

  $f(x)=P(X=x)$
- olarak göstereceğiz.

- Bu örnek için
- $f(0)=P(X=0)=1/4$(hiç tura olamayan bir tane sonuç vardır: YY)
- $f(1)=P(X=1)=2/4$(bir tura olan sonuç sayısı iki; YT,TY)
- $f(2)=P(X=2)=1/4$(ikisi de tura olan bir tane sonuç vardır: TT)
- Daha özet bir gösterim ile

$X=x$	0	1	2
$P(X=x)$	¼	½	¼

- yazılır. Burada dikkat edilirse, $P(X=0)+P(X=1)+P(X=2)=1/4+1/2+1/4=1$ dir.
- Olasılık dağılımı için X, sonlu sayıda x1,x2,...,xn değerlerini
- $f(x_i)=P(X=x)$ $i=1,2,...,n$

- olasılıkları ile alabilen kesikli raslantı değişkeni olsun.
- $P(X=x_i)=f(x_i)$ aşağıda verilen koşulları sağlıyor ise X'in olasılık dağılımı ya da olasılık fonksiyonudur. Bu koşullar :
- $P(X=x_i) \geq 0$ tüm x_i'ler için
- $\sum_{i=1}^{n} P(X = x_i) = 1$ dir.
- Örneklem uzayı sonsuz ise ikinci koşul,
- $\sum_{i=1}^{\infty} P(X = x_i) = 1$ dur.